

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Application No. :

U.S. National Serial No. :

Filed :

PCT International Application No. : PCT/DE03/03552

VERIFICATION OF A TRANSLATION

I, the below named translator, hereby declare that:

My name and post office address are as stated below;

That I am knowledgeable in the German language in which the below identified international application was filed, and that, to the best of my knowledge and belief, the English translation of the international application No. PCT/DE03/03552 is a true and complete translation of the above identified international application as filed.

I hereby declare that all the statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the patent application issued thereon.

Date: March 16, 2005

Full name of the translator : Michael Amos WHITTAKER

For and on behalf of RWS Group Ltd

Post Office Address : Europa House, Marsham Way,  
Gerrards Cross, Buckinghamshire,  
England.

## Description

Method for fabricating a transistor structure

5

The present invention relates to a method for fabricating a transistor structure, comprising at least a first and a second bipolar transistor having different collector widths. Such a method is known from  
10 DE 100 44 838 C2 for example.

In bipolar transistors, the collector is usually terminated by a highly doped buried layer. The buried layer is produced by subjecting the substrate to an ion 15 implantation at the desired location. Afterward, a lightly doped epitaxial layer is applied and the wells for base, emitter and collector are produced. A possible process sequence is described for example in the textbook "Technologie hochintegrierter Schaltungen"  
20 ["Technology of large scale integrated circuits"] by D. Widmann, H. Mader, H. Friedrich, Springer Verlag, 2nd edition, Table 8.13, pp. 326-334.

For integrated high-frequency circuits in the GHz 25 range, it is favorable to integrate both high-voltage transistors (HV transistors) having a high breakdown voltage and high-frequency transistors (HF transistors) having a short collector transit time and hence a high limiting frequency  $F_L$ . On account of the fabrication 30 methods known heretofore, it is necessary to find a compromise with regard to the properties in the integration of bipolar transistors having different limiting frequencies and bipolar transistors having different breakdown voltages in high-frequency 35 circuits. This means that the performance of such a high-frequency circuit cannot be utilized optimally.

Such integration has been realized hitherto e.g. by means of the collector region having a dopant

concentration of differing magnitude. The lower the doping, the higher the collector-base breakdown voltage. However, as a result of this the collector transit time becomes longer and hence the limiting frequency  $F_t$  becomes lower. The higher the doping, the shorter the collector transit time but the smaller the collector-base breakdown voltage of the transistor.

10 M. Racanelli et al. "Ultra High Speed SiGe NPN for Advanced BiCMOS Technology", 2001 IEEE discloses scaling the doping of a collector region of a transistor in such a way that the dopant concentration within the collector region has a gradient. Although this solution makes it possible to increase the  
15 breakdown voltage of HF transistors, it still represents a compromise.

In addition to the dopant concentration, the dimensioning of the collector width also determines the properties of a bipolar transistor. The term collector width denotes that region of the epitaxial layer which is located between the base well located in the epitaxial layer and the buried layer. HF transistors which are to be optimized toward high limiting  
25 frequencies must have a small collector width and HV transistors which are optimized toward high breakdown voltages must have a large collector width.

30 DE 100 44 838 C2 describes a semiconductor component and also a method for fabricating it in which bipolar components having different collector widths are realized. In this case, an additional substance is introduced into a buried layer of a bipolar component, which additional substance influences the diffusion of  
35 a dopant of the buried layer and hence the collector width of said bipolar component. However, this method does not result in a sharp junction between the differently doped buried layers and collectors.

Consequently, the collector width cannot be established exactly and with a sharp profile, but rather exhibits a "blurred" profile with a shallow gradient.

5 Accordingly, it is an object of the present invention to specify an optimized method for fabricating a transistor structure in which collector regions having different collector widths can be formed, the collector regions having a sharp boundary with respect to the  
10 buried layers.

According to the invention, the object is achieved by means of a method of the type mentioned in the introduction in which there are produced at least a  
15 first collector region having a first collector width C1 on a first buried layer and a second collector region having a second collector width C2 on a second buried layer, in which case, for the production of the second collector width C2, a first collector zone  
20 having a first thickness C3 is produced on the second buried layer and a second collector zone having a second thickness C4 is produced on the first collector zone, and at least one insulation region is produced, which isolates at least the collector regions from one  
25 another.

What is achieved as a result of this is that the two bipolar transistors of the transistor structure have a different collector width, and the collector regions  
30 have a sharp or abrupt junction with a steep gradient with respect to the adjacent regions, such as the buried layers. The collector width C1 of the first bipolar transistor preferably corresponds to the first thickness C3 of the second collector region. The  
35 collector width C2 of the second bipolar transistor is composed of the thicknesses C3 and C4 of the collector zones of the second collector region. Accordingly, the thicker the second thickness C4, the greater also the

difference between the collector widths of the two bipolar transistors.

The invention is based on the insight that a sharp  
5 boundary or an abrupt junction between the lightly doped collector and the highly doped buried layer significantly improves the behavior of a transistor since an abrupt profile with a steep gradient, with a sheet resistance remaining the same, has lower fringing  
10 capacitances than a profile with a shallow gradient. Equally, the method according to the invention improves the high-current behavior of the transistor since there is no unnecessary dopant in that part of the collector which is flooded with charge carriers and, by the same  
15 token, the conduction of the buried layer is lowered.

According to the invention, the method of the type mentioned in the introduction is developed further to the effect that at least a first zone of a first buried layer of a first conductivity type of the first bipolar transistor and a first zone of a second buried layer of a first or a second conductivity type of the second bipolar transistor are introduced into the semiconductor substrate, a first epitaxial layer is  
20 produced, which covers, over the whole area, at least the first zone of the buried layers, at least a second zone of the first conductivity type is produced within the first epitaxial layer, the second zone adjoining the first zone of the first buried layer, a second  
25 epitaxial layer is produced, which covers, over the whole area, at least the first epitaxial layer and the second zone of the first buried layer, at least one insulation region is produced, which isolates at least the collector regions from one another, the second zone  
30 of the first buried layer adjoining the first collector region and the first zone of the second buried layer adjoining the second collector region.  
35

Buried layers having different thicknesses are thereby produced, the thickness of the first buried layer being composed of a first zone, which is introduced into the semiconductor substrate, and a second zone, which is  
5 introduced into the first epitaxial layer. The second buried layer and the first zone of the first buried layer preferably have the same thickness in this case. The thicknesses of the first and second buried layers thus differ by the thickness of the second zone of the  
10 second buried layer. Since the collector width, as mentioned in the introduction, depends on the thickness of the epitaxial layer, minus the buried layer extending into the epitaxial layer, the collector widths C1 and C2 can be varied in a simple manner and,  
15 unlike in embodiments known hitherto, nevertheless have a sharp junction between the highly doped buried layers and the lightly doped collector regions.

The inventors furthermore propose developing the method  
20 of the type mentioned in the introduction further in such a way that at least a first zone of a first buried layer of a first conductivity type of the first bipolar transistor and a second buried layer of a first or a second conductivity type of the second bipolar  
25 transistor are introduced into the semiconductor substrate, at least a first collector zone of the first bipolar transistor and a first collector zone of the second bipolar transistor are produced, the first collector zone of the first bipolar transistor adjoining the first zone and the first collector zone of the second bipolar transistor adjoining the second buried layer, the first collector zone is formed as first conductivity type, a second collector zone is produced on the first collector zone of the second bipolar transistor and a second collector zone is produced on the first collector zone of the first bipolar transistor, and at least one insulation region  
30  
35

is produced, which isolates at least the collector zones from one another.

This too means that it is possible, in a simple manner,  
5 for collector regions to be fabricated which have both different thicknesses and sharp profiles with a steep gradient with respect to the adjacent buried layers and, consequently, a transistor structure is formed which comprises both the properties of an HV transistor  
10 and those of an HF transistor.

A development of the methods according to the invention provides for the third collector zone to be deposited.

15 In a preferred development of the methods according to the invention, the third collector zone is deposited epitaxially. As a result of this, the collector zone is grown with the fewest possible crystal defects, which is very important for the functional properties of a  
20 bipolar transistor.

Another refinement provides for an insulating layer (SOI layer, SOI - Silicon On Insulator) to be provided between the buried layers and the semiconductor substrate.  
25 The collector regions are thereby electrically insulated and capacitively decoupled from the substrate without the need for any additional insulation.

30 Typically, the insulation region which isolates at least the collector regions from one another is isolated from one another with the aid of shallow trench isolation technology (STI technology). The insulation region may be filled with an electrically  
35 insulating material, such as a CVD oxide (CVD = Chemical Vapor Deposition), for example. Preferably, the two laterally adjacent highly doped buried layers of two bipolar transistors are thereby electrically

insulated from one another. Said insulation region may be embodied for example as a full trench or as a deep trench.

5 A full trench is a trench, for example between components of a chip, in which the silicon is etched or interrupted as far as the buried layers, so that current paths between the components are completely interrupted. A full trench can isolate relatively large  
10 transistor regions from one another, as is also described in an article by S. Maeda, "Impact of 0.18  $\mu\text{m}$  SOI CMOS Technology using Hybrid Trench Isolation with High Resistivity Substrate on Embedded RF/Analog Applications", 2000 Symp. on VLSI Technology - Digest  
15 of Technical Papers (CAT. No. 00CH37104), pages 154 to 155.

A deep trench is described for example in the article "An SOI-Based High Performance Self-Aligned Bipolar  
20 Technology Featuring 20 ps Gate-Delay and a 8.6 fJ Power Delay Product" by E. Bertagnolli et al., 1993, Symp. on VLSI Technology, Digest of Technical Papers (CAT. No. 93CH3303-5), pages 63 to 64. In contrast to the full trench, the deep trench is not wide enough to  
25 be able to integrate the entire dimensions of passive components above it. Rather, the deep trench serves for dielectric component isolation.

Preferred exemplary embodiments of the present  
30 invention are explained in detail below with reference to the accompanying drawings, in which

35 figures 1a to 1d show a diagrammatic cross-sectional view of a first method according to the invention for producing a transistor structure according to the invention with two collector regions having a different

collector width by means of selective epitaxy,

figures 2a to 2e show a diagrammatic cross-sectional view of a second method according to the invention for producing a transistor structure with two collector regions having a different collector width by means of whole-area epitaxy,  
5

figures 3a to 3c show a diagrammatic cross-sectional view of a third method according to the invention for producing a transistor structure with two collector regions having a different collector width,  
10  
15

figure 4 shows a diagrammatic cross-sectional view of an alternative configuration for producing a transistor structure with two collector regions having a different collector width with an SOI structure.  
20  
25

The first method according to the invention for producing a transistor structure according to the invention with two collector regions having a different collector width which is described below with reference to figures 1a to 1d is performed by means of selective epitaxy.  
30

In figure 1a, the, for example n<sup>+</sup>-doped, buried layers 5.1 and 5.2 have already been introduced into the semiconductor substrate 1 and insulated from one another by insulation regions 4, here realized as deep  
35

trenches 4. The semiconductor substrate 1 comprises e.g. monocrystalline silicon which is p-doped.

5 A first auxiliary layer 6 and a second auxiliary layer  
7 are furthermore provided, which isolate the  
insulation region 4 from the semiconductor substrate 1  
and the buried layers 5.1 and 5.2. In this case, the  
second auxiliary layer 7 adjoins the insulation region  
4 and the first auxiliary layer 6 adjoins the second  
auxiliary layer 7 and also the semiconductor substrate  
10 1 and the buried layers 5.1 and 5.2.

15 The second auxiliary layer 7 preferably comprises a  
material which is resistant to oxidation and can be  
etched selectively with respect to oxide, such as  
silicon nitride  $\text{Si}_3\text{N}_4$ , for example. This makes it  
possible to avoid sidewall defects, that is to say  
crystal defects which arise at the interface between a  
dielectric material and silicon during the epitaxial  
20 growth of the silicon. In another variant, the second  
auxiliary layer 7 may be formed from polysilicon. The  
thickness of said second auxiliary layer 7 lies in the  
range of between 3 nm and 60 nm. One advantage of this  
thin nitride lining is that it protects the wall of the  
25 insulation region 4, which is filled with CVD oxide,  
for example, against oxidations and thus prevents  
defect formations.

30 The first auxiliary layer 6 preferably comprises a  
material which can be etched selectively with respect  
to the layer 7 and avoids large mechanical stresses on  
the semiconductor substrate 1, such as an oxide, for  
example. Furthermore, the auxiliary layer 6, up to the  
epitaxy, can protect the sensitive silicon surface of  
35 the buried layers 5.1 and 5.2 against oxidation.

The openings 12 are etched into an STI oxide layer 13  
down to the second auxiliary layer 7, which STI oxide

layer preferably covers the whole area of the semiconductor substrate 1 with the auxiliary layers 6 and 7 situated thereon. As is known from EP 0 600 276 B1, the etching may be effected by 5 anisotropic dry etching which stops selectively on silicon nitride, and thus on the second auxiliary layer 7.

In the subsequent method step in accordance with figure 10 1b, a lateral undercut 14 of the auxiliary layers 6 and 7 is effected. The undercut 14 is described in more detail in EP 0 600 276 B1. Since sidewall defects form proceeding from the interfaces between the auxiliary layers 6 and 7 and the surface of the buried layers 5.1 15 and 5.2 and grow up at an angle of about 52° along (111) crystal faces, that is to say for example along the sidewall of the STI oxide layer, this growth of the sidewall defects can be interrupted by the overhang formed by the undercuts 14 of the STI oxide layer 13.

20 Afterward, a first collector region 2.1 having a thickness C1 and also a collector zone 2.2.1 having a thickness C3 are deposited epitaxially, the first collector region 2.1 adjoining the first buried layer 25 5.1 and the collector zone 2.2.1 adjoining the second buried layer 5.2. In this case, the collector zone 2.2.1 is provided for the second collector region of a second bipolar transistor. The thicknesses C1 and C3 of the collector region 2.1 and of the collector zone 30 2.2.1 are approximately identical and are preferably between 50 nm and 300 nm.

After the first collector region 2.1 has been covered with a masking layer 8, a further collector zone 2.2.2 35 is applied epitaxially on the collector zone 2.2.1 in figure 1c. Said collector zone 2.2.2 preferably has a thickness C4 of between 100 nm and 200 nm. The collector width C2 of the second collector region 2.2,

composed of the collector zones 2.2.1 and 2.2.2, thus lies in the range of between 150 nm and 500 nm.

5 In the present example, the second collector region 2.2 terminates at approximately the same level with the surface of the STI oxide layer 13.

10 Typically, the collector width C1 of the first collector region 2.1 and the collector width C2 of the second collector region 2.2 are in a ratio of between 0.05 and 0.9 to one another. Typical values are 100 nm for the collector width C1 and 250 nm for the collector width C2. The different collector widths C1 and C2 of the two collector regions 2.1 and 2.2 on the same 15 semiconductor substrate 1 have the effect of optimizing the properties both of an HF transistor and of an HV transistor.

20 If the intention is to obtain even greater differences between the collector widths of the first and second collector regions 2.1 and 2.2, then, in the method step of figure 1b, the collector width C1 and the thickness C3 are kept relatively low and, in the subsequent method step of figure 1c, the deposition of the 25 collector zone 2.2.2 with a second thickness C4 is repeated appropriately often.

30 In the transistor structure of figure 1d, the masking layer above the collector region 2.1 has been removed and collector terminal regions 11 have been introduced. After filling with tungsten, for example, the collectors can be routed out electrically to the surface, thereby enabling the transistor structure to be integrated into an integrated circuit.

35

The method according to the invention for fabricating a transistor structure for two bipolar transistors with the aid of selective epitaxy which has been described

with reference to figures 1a to 1d is particularly simple. The various collector regions are deposited in respectively successive epitaxy steps with the required thickness in the STI oxide layer 13, the already 5 completed collector region 2.1 being covered by a masking layer 8 in order to prevent a further epitaxial deposition. The regions in the STI oxide layer 13 which are required for an epitaxial deposition are thus opened in each case only for the corresponding epitaxy 10 step.

It is also possible with the aid of whole-area epitaxy to realize collector regions 2.x having different collector widths and sharp junctions with respect to 15 the highly doped buried layers, as explained below with reference to figures 2a to 2e. In this case, the collector regions 2.x terminate toward the top in planar fashion at the same level with the surface of the STI oxide layer 13, the thicknesses D1 and D2 of 20 the buried layers 5.1 and 5.2 being varied. This planar termination is particularly advantageous since planar surfaces are required for subsequent photosteps (not described here), with feature sizes of less than 0.35  $\mu\text{m}$ .

25 In accordance with figure 2a, an epitaxial layer 9 having a thickness E1 is deposited over the whole area of the semiconductor substrate 1, into which a first zone 5.1.1 of a first buried layer and a further first 30 zone 5.2.1 of a second buried layer have already been implanted. The first zones 5.1.1 and 5.2.1 are preferably n<sup>+</sup>-doped.

Afterward, in figure 2b, a second zone 5.1.2 of the 35 buried layer 5.1 and a second zone 5.2.2 of the buried layer 5.2 are introduced into the epitaxial layer 9, these second zones 5.x.2 also being n<sup>+</sup>-doped. In this case, the second zone 5.1.2 extends approximately over

the area of the first zone 5.1.1, whereas the second zone 5.2.2 of the buried layer 5.2 merely extends over a partial region of the first zone 5.2.1 of the buried layer 5.2.

5

In the subsequent step of the method according to the invention in figure 2c, a second epitaxial layer 10 is deposited with a thickness E2 over the whole area of the epitaxial layer 9 and the second zones 5.x.2 of the 10 buried layers 5.1 and 5.2. In this case, said second epitaxial layer 10 may result from a single deposition or a plurality of depositions in succession. The collector width C1 of a first collector region may be defined by the thickness E2 of said epitaxial layer 10, 15 the collector width C1 corresponding to the thickness E2 of the epitaxial layer 10. By contrast, the collector width C2 of the second collector region corresponds to the sum of the thicknesses E1 of the epitaxial layer 9 and the thickness E2 of the epitaxial 20 layer 10.

In figure 2d, the buried layers 5.1 and 5.2 are insulated from one another by means of insulation regions 4, embodied as deep trenches in the present 25 example.

Afterward, in figure 2e, the STI oxide layer 13 is etched into the epitaxial layer 10 in accordance with figure 2d and preferably filled with STI oxide, regions 30 for the collector terminal region 11 and also the first and second collector regions 2.1 and 2.2 being left free. The collector terminal regions 11 are then etched above the second zones 5.2.2 and 5.1.2 in order thereby to enable an electrical linking of the collector.

35

The first collector region 2.1 thus has a first collector width C1 and the second collector region has a larger collector width C2. The two collector regions

2.1 and 2.2 terminate in planar fashion with the surface of the STI oxide layer 13 and both have sharp junctions between the heavily doped regions of the buried layers 5.x and the more lightly doped collector 5 regions 2.x. As a result of this, the transistor structure acquires defined and exactly determinable properties.

A further method according to the invention for 10 fabricating a transistor structure according to the invention with two collector regions having a different collector width is described in more detail with reference to figures 3a to 3c.

15 In figure 3a, analogously to figure 1b, provision is made of a structure having a, preferably p-doped, semiconductor substrate 1, a first zone 5.1.1 of a first buried layer implanted into the semiconductor substrate 1, and an implanted second buried layer 5.2, 20 having a thickness D2, insulation regions 4, a first auxiliary layer 6 and a second auxiliary layer 7, an STI oxide layer 13 and the collector zones 2.1.1 and 2.2.1.

25 As in figure 1b, in figure 3a, the auxiliary layers 6 and 7 are undercut under the STI oxide layer 13, so that the collector zones 2.1.1 and 2.2.1 have a stepped profile in cross section. By virtue of this undercut 14, the STI oxide layer 13 exhibits an overhang over 30 part of the collector zones 2.1.1 and 2.2.1.

The thickness C3 of said collector zones 2.1.1 and 2.2.1 may vary between 5 nm and 300 nm.

35 After a masking layer 8 has been applied to the region of the collector zone 2.2.1, the collector zone 2.1.1, as shown using the arrows 15 in figure 3b, is doped in such a way that it has the same doping as the first

zone 5.1.1 of the first buried layer 5.1. This is preferably an n<sup>+</sup>-type doping. This newly formed second zone 5.1.2 and the first zone 5.1.1 now form the buried layer 5.1 having a thickness D1.

5

After the removal of the masking layer 8, in figure 3c, a first collector region 2.1 is deposited epitaxially on the first buried layer 5.1 with a collector width C1 and a further collector zone 2.2.2 having a thickness 10 C4 is deposited epitaxially on the collector zone 2.2.1. The second collector region 2.2 is now formed from the two collector zones 2.2.1 and 2.2.2 and has a collector width C2 representing the sum of the thicknesses C3 and C4. Both collector regions 2.1 and 15 2.2 terminate in planar fashion with the surface of the STI oxide layer 13.

After the introduction of the collector terminal regions 11 and filling with tungsten, for example, the 20 transistor structure of figure 3c is suitable for use in bipolar transistors.

In a further embodiment according to the invention as shown in figure 4, an insulating layer 3 is produced 25 between the semiconductor substrate 1 and the buried layers 5.1 and 5.2.

The production of the collector regions 2.1 and 2.2 having different collector widths C1 and C2 and sharp 30 junctions between the collector regions and the buried layers may correspond to the method according to the invention from figures 3a to 3c. Moreover, the methods according to the invention which have been explained with reference to figures 1a to 1d and 2a to 2e are 35 also conceivable.

Preferably, in the methods according to the invention as shown in figures 1 to 4, the first buried layer 5.1

and the second buried layer 5.2 are configured as an identical conductivity type. This configuration makes it possible to form two transistor structures of identical type one beside the other, that is to say for 5 example two npn transistors or two pnp transistors.

In an alternative embodiment, the first buried layer 5.1 and the second buried layer 5.2 are configured as different conductivity types. This makes it possible to 10 form an npn transistor beside a pnp transistor on the same semiconductor substrate 1.

In a particularly advantageous development of the methods according to the invention, the collector 15 regions are formed with a dopant gradient, the concentration of the dopant varying in the horizontal direction. This development makes it possible, for example, to form an increased dopant concentration in the central collector region. This development, in 20 particular with a small thickness of the collector region, reduces the base-collector space charge zone and thus reduces the collector transit time. This development is particularly advantageous primarily in the case of very small transistor structures in which 25 the emitter region is arranged centered above the collector region.

It goes without saying that, instead of the deep trenches, it is also possible to use full trenches as 30 insulation regions 4.

Overall, the methods according to the invention as explained with reference to figures 1 to 4 make it possible to fabricate transistor structures having a 35 first collector region 2.1, having a first collector width C1, and also a second collector region 2.2, having a larger collector width C2, on the same semiconductor substrate 1, all junctions between

differently doped regions having a sharp interface. In this case, by way of example, the first collector region 2.1 is suitable for a high-frequency transistor with high limiting frequencies  $f_T$  and the second 5 collector region 2.2 is suitable for a high-voltage transistor with increased breakdown voltages.